DOCUMENT RESUME

ED 310 931 SE 050 875

AUTHOR Mestre, Jose; And Others
TITLE Hierarchical Problem Solving as a Means of Promoting Expertise.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 88
GRANT BNS-8511069
NOTE 9p.; Small, broken type in Figure 1 may not reproduce well.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Processes; *College Science; *Computer Assisted Instruction; Computer Oriented Programs; Computer Uses in Education; Decision Making Skills; Higher Education; *Physics; *Problem Solving; *Science Instruction; Undergraduate Study

ABSTRACT

This paper reports on the effects of structuring problem solving activities of novices in the domain of physics in a way consistent with the problem solving approaches used by experts. A total of 42 undergraduate students who completed a first semester physics course and received a grade of B or better participated in this study. The treatment involved five 1-hour sessions during which subjects solved a total of 25 classical mechanics problems using a hierarchical, computer-based, problem-analysis environment called the Hierarchical Analysis Tool (HAT). Two types of tasks were administered before and after treatment: (1) a problem-categorization task; and (2) a problem-solving task. The results indicate that the hierarchical approach to problem solving as exemplified by the HAT helped students to shift their decision-making criteria for problem categorization from one based on surface features toward one based on deep structure. Two implications from this study were discussed. Twelve references are listed. (YP)

HIERARCHICAL PROBLEM SOLVING AS A MEANS OF PROMOTING EXPERTISE

Jose Mestre, Robert Dufresne

William Gerace, Pamela T. Hardiman

Scientific Reasoning Research Institute
University of Massachusetts at Amherst

1988

Hierarchical Problem Solving as a Means of Promoting Expertise*

Jose Mestre, Robert Dufresne, William Gerace and Pamela T. Hardiman
University of Massachusetts at Amherst

INTRODUCTION

Becoming an expert in a given domain takes substantial time and effort. This raises an interesting question: For a given individual possessing aptitude in some domain, is becoming an expert simply a function of time and effort, or can the path toward expertise be made more efficient? The research reported in this article reports on the effects of structuring the problem solving activities of novices in the domain of physics in a way consistent with the problem solving approaches used by experts. The focus of our investigation was to assess the possibility of promoting expert-like behavior among novices by constraining them to follow an expert-like approach to problem analysis.

That novices and experts store and use domain-specific knowledge in distinctly different ways is the consensus of a number of studies in such diverse fields as chess (Chase & Simon, 1973), computer programming (Ehrlich & Soloway, 1982), electrical circuits (Egan & Schwartz, 1979), and classical mechanics (Larkin, 1979). Experts tend to store information in hierarchically structured clusters related by underlying principles or concepts. When attempting to solve a problem, experts initially focus on the principles and heuristics that could be applied to solve that problem (referred to as deep structure cuing). In contrast, the knowledge base of novices is less structured and has fewer interconnections. When solving problems, novices do not focus on principles or heuristics that could be used to construct a solution strategy; rather, they focus on objects and descriptor terms in the problem (called surface features) and then look for the actual equations that could be manipulated to yield an answer (Chi, Feltovich & Glaser, 1981; Larkin, McDermott, Simon & Simon, 1980; Mestre & Gerace, 1986).

Studies in the domain of physics (Eylon & Reif, 1984; Heller & Reif, 1984) suggest that instructional approaches that impose a hierarchical, expert-like organization both on information, and on problem solving heuristics result in improved problem solving and recall performance among novices. Despite these, and other related findings current instructional practice does not emphasize hierarchical approaches to knowledge organization or to problem solving. Consequently, much of the the expert's tacit knowledge remains a secret to the novice until she or he discovers it on her own.

In the present study novices actively participated in problem solving activities which were structured to reflect our best understanding of how physics experts analyze problems. The treatment involved five one-hour sessions during which subjects solved a total of 25 classical mechanics problems using a hierarchical, computer-based, problem-analysis environment called the Hierarchical Analysis Tool. The effectiveness of this treatment

---

* Supported by National Science Foundation grant BNS-8511069. The contents herein do not necessarily reflect the position, policy or endorsement of NSF.

was compared with that of two control treatments in which novices solved the same problems using more traditional, novice-like approaches. To assess the effectiveness of the treatments at promoting shifts toward expertise, two types of tasks were administered before and after treatment: a problem categorization task (discussed in Experiment 1) and a problem solving task (discussed in Experiment 2).

DESCRIPTION OF COMPUTER-BASED ENVIRONMENTS USED IN TREATMENTS

Appreciating the findings of this study requires some understanding of the treatments that subjects received. In this section we briefly describe the architecture and functioning of the Hierarchical Analysis Tool (henceforth HAT) used in the focal treatment. We also describe another computer-based, equation-based environment used in one of the two control treatments.

The HAT is a menu-driven environment that constrains users to perform a top-down analysis of classical mechanics problems. It combines declarative and procedural information in a hierarchical framework. The environment is capable of handling the majority of problems in a typical calculus-based freshman-level classical mechanics course. The word "tool" in the name implies that the HAT can be used to facilitate the construction of a solution; the HAT does not supply answers to problems.

To analyze a problem, the user answers well-defined questions by making selections from menus that are dynamically generated by software. In the first menu, the user selects one of four general principles that could be applied to solve the problem under consideration. Subsequent menus focus on ancillary concepts and procedures, and are dependent upon the prior selections made by the user. When the analysis is complete, the HAT provides the user with a set of equations that is consistent with the menu selections made during the analysis. If the analysis is carried out appropriately, then these equations could be used to generate a solution to the problem; however, the user must still manipulate these equations to isolate the quantity asked for in the problem. In the event that the analysis is carried out incorrectly, the final equations are consistent with the user's choices, but inappropriate for solving the problem. It is important to note that the HAT neither tutors nor provides feedback to the user--it merely constrains the type and order of questions that need be considered when analyzing a problem. Figure 1 provides a sample problem and the HAT menus and selections that would appropriately analyze the problem.

A second computer-based environment, called the Equation Sorting Tool (EST), was developed for use as a control treatment. The EST was designed to be consonant with the approach taken by most novice physics students. It is a data base of 178 equations taken from a standard classical mechanics textbook. This equation data base can be sorted in three different ways: 1) by Problem Types, such as "inclined plane" and "falling bodies," 2) by Variable Names, such as "mass" and "velocity," and 3) by Physics Terms, such as "potential energy" and "momentum." By "sorted" we mean that the user can perform sequential logical "ands" to narrow down the data base to a small, manageable number of equations that might be useful for solving a problem. The EST was designed to reflect novices' problem solving tendencies in physics: they tend to cue on surface features in deciding how to attack a problem, and focus their problem solving efforts on finding the appropriate set of equations that can be manipulated to yield an answer to the problem. Further details on the design and functioning of both the HAT and the EST can be found elsewhere (Dufresne, Gerace, Hardiman & Mestre, 1987; Mestre & Gerace, 1986).

FIGURE 1. Hierarchical Analyzer Menus & Choices for Problem 1

1 Which principle applies to this part of the problem solution?

1. Newton's Second Law or Kinematics
2. Angular Momentum
3. Linear Momentum
4. Work and Energy

Please enter your selection: [4]

(B)ackup (M)ain menu (G)lossary (Q)uit (L)ist selections

2 Describe the system in terms of its mechanical energy

1. Conservative system (conservation of energy)
2. Non-Conservative system (work-energy exchange)

Please enter your selection: [1]

(B)ackup (M)ain menu (G)lossary (Q)uit (L)ist selections

3 Describe the changes in mechanical energy. Consider only the energy of one body at some initial and final state

1. Change in kinetic energy
2. Change in potential energy
3. Change in potential and kinetic energies

Please enter your selection: [3]

(B)ackup (M)ain menu (G)lossary (Q)uit (L)ist selections

4 Describe the changes in kinetic energy

1. Change in translational kinetic energy
2. Change in rotational kinetic energy
3. Change in translational and rotational kinetic energies

Please enter your selection: [1]

(B)ackup (M)ain menu (G)lossary (Q)uit (L)ist selections

5 Describe the boundary conditions

1. No initial translational kinetic energy
2. No final translational kinetic energy
3. Initial and final translational kinetic energies

Please enter your selection: [1]

(B)ackup (M)ain menu (G)lossary (Q)uit (L)ist selections

6 Describe the changes in potential energy

1. Changes in gravitational potential energy
2. Changes in spring potential energy
3. Changes in gravitational and spring potential energies

Please enter your selection: [1]

(B)ackup (M)ain menu (G)lossary (Q)uit (L)ist selections

7 Describe the boundary conditions

1. No initial gravitational potential energy
2. No final gravitational energy
3. Initial and final gravitational energy

Please enter your selection: [2]

(B)ackup (M)ain menu (G)lossary (Q)uit (L)ist selections

8 Is there another body in the system which has not been examined?

1. Yes
2. No

Please enter your selection: [2]

(B)ackup (M)ain menu (G)lossary (Q)uit (L)ist selections

9 The Energy Principle states that the work done on the system by all non-conservative forces is equal to the change in the mechanical energy of the system:

$$W_{nc} = E_f - E_i$$

According to <u>your</u> selections,

$W_{nc} = 0$ (Conservative system: mechanical energy conserved)

$E_f = (\frac{1}{2}Mv^2)_{1f}$

$E_i = (Mgy)_{1i}$

Please press any key to continue

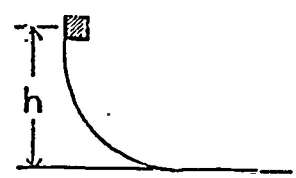
10 *** Work and Energy ***

1. Problem solved
2. Return to Main Menu to continue solution
3. Review previous solution screens

Please enter your selection:

PROBLEM 1

A SMALL BLOCK OF MASS M SLIDES ALONG A TRACK HAVING BOTH CURED AND HORIZONTAL SECTIONS AS SHOWN. IF THE PARTICLE IS RELEASED FROM REST AT HEIGHT h, WHAT IS ITS SPEED WHEN IT IS ON THE HORIZONTAL SECTION OF THE TRACK? THE TRACK IS FRICTIONLESS.



PARTICIPANTS IN EXPERIMENTS 1 AND 2

Subjects

Forty-two undergraduate students at the University of Massachusetts who had completed the first semester physics course for majors or for engineers, and received a grade of B or better, participated in this study. The subjects participated in ten hour-long experimental sessions, for which they were paid fifty dollars.

Groups

On the basis of pretest scores, the 42 subjects were divided into three treatment groups of 14 subjects each. Each of the three groups received the same 25 problems over the course of the treatment. Subjects solved five problems in each of five sessions over approximately three weeks; the treatment problems were representative of problems the subjects had encountered in their course and covered the major topics in a beginning classical mechanics course. The "HAT-group" used the Hierarchical Analysis Tool in solving the 25 treatment problems. The "EST-group" used the Equation Sorting Tool in solving the treatment problems, while the "T-group" used a homework-style approach in solving the treatment problems and were free to refer to the textbook that they had used in their course to solve the problems.

EXPERIMENT 1: SIMILARITY JUDGMENT TASK

We designed a similarity judgment task (Hardiman, Dufresne & Mestre, 1987) in which subjects were to decide which of two comparison problems would be solved most similarly to a third model problem. Surface feature and deep structure similarity to the model problem were varied systematically, allowing us to investigate whether subjects were more likely to focus on deep structure similarity as a basis for categorization following treatment. Since the initial decision that must be made in the HAT concerns the principle to be applied, we hypothesized that the HAT-group would be more likely to focus on deep structure after treatment than either of the two control groups.

This task contained 20 items. Each task item was composed of three elementary mechanics problems, each of which was three to five lines long and contained only text (no pictures or diagrams). For each item, one of the three problems was identified as the <u>model problem</u>, while the other two were the <u>comparison problems</u>. The subjects were to indicate which of the two comparison problems they believed "would be solved most similarly" to the model problem.

A comparison problem could share different attributes with its model problem. Four types of comparison problems were designed that matched the model problem in: 1) surface features, meaning that the objects and descriptor terms that occur in both problems are similar, 2) deep structure, meaning that the physical principle that could be applied to solve both problems is the same, 3) both surface features and deep structure, or 4) neither surface features nor deep structure. These four types of comparison problems were termed S, D, SD, and N, respectively.

The comparison problems were paired such that only one of the two comparison problems matched the model problem in deep structure. This constraint led to four types of comparison problem pairs: 1) S-D, 2) S-SD, 3) N-D, and 4) N-SD. Assuming a categorization scheme based strictly on surface features, the following pattern of performance was predicted: 1) S-D: 0% deep structure choices, 2) S-SD: 50% deep structure choices (both choices are equally good in terms of matching the model problem on surface features), 3) N-D: 50% deep structure choices (either alternative is equally "bad" in terms of matching the model problem on surface features), and 4) N-SD: 100% deep structure choices (a surface feature match to the model problem will also mean a deep structure match). In contrast, assuming a deep structure categorization scheme would result in 100% deep structure choices in all four pairings.

The task was presented via computer. The subject was told to read carefully the model problem and two comparison problems, and to respond by pressing one of two keys. The items were presented in random order, with no limit imposed on time to respond. After every 5 items, the subject was given the opportunity to take a brief rest. Most subjects completed the task within 45 minutes. The same task was presented after the subjects had completed the five treatment sessions.

Results and Discussion

The performances of the 42 subjects were compared in a 3 (Treatment Groups) x 14 (Subjects/Group) x 2 (Times: pre & post) x 4 (Comparison Problem Pairings) x 5 (Model Problems) analysis of variance. The focal question was whether the HAT treatment would promote a shift toward reliance on deep structure rather than on surface features. The results indicate an affirmative response to this question. In the pre-judgment task, there were no differences between the groups, as can be seen in Table 1. However, there were significant differences among the three groups in the amount of improvement from pre- to post-treatment on the judgment task (see Table 1), $F(2,39) = 4.28$, $p = .02$.

The HAT-group was the only group to show any indications of improvement; their improvement was statistically significant, ($F(1,13) = 5.20$, $p = .04$). In contrast, the mean performance of the EST-group remained the same, while the performance of the T-group declined. This result suggests that the HAT does promote a shift toward the use of deep structure, while the two control treatments do not. This shift was consistent across Comparison Problem Types, with improvements of at least 6 percentage points in each of the four comparison problem pairings. This improvement was significant for the S-D pairings (a 17% pre-to-post improvement), $t(13) = 3.12$, $p = .0324$ (adjusted for four tests), which is encouraging given that the S-D items, where surface features and deep structure are in direct competition, present the most difficulty for novices (Hardiman, et al., 1987).

Table 1: Pre- and Post-Judgment Task Percent Correct for the 3 Groups

Group	pre-treatment	post-treatment
HAT	56%	66%
EST	61%	61%
T	62%	58%
Total	60%	62%

EXPERIMENT 2: PROBLEM SOLVING TASK

To measure the effect of treatment on problem solving, two equivalent tests were constructed in the style of a traditional final exam for a freshman level classical mechanics course. Half of the subjects received one form of the test on the pre-assessment while the other half received the second form. The form not used on the pre-assessment was used for the post-assessment. The tests contained 7 problems, with 4 questions requiring the application of one physical principle for solution and the remaining 3 problems requiring the application of two principles. Subjects were given approximately one hour to solve all 7 problems.

We expected that all subjects would improve in performance from the pre- to the post-test, since all subjects would have practiced solving problems during the treatment phase. If the HAT-group was capable of adopting and applying the concept-based approach on the post-test, we might expect them to exhibit better pre-to-post improvements than the EST- and T-groups.

Results and Discussion

The tests were graded independently by two physicists. Whenever the score on an item differed between the graders, the subject's solution was reevaluated and a score was determined by consensus. The pre- and post-test scores are shown in Table 2. All three groups increased about 10 percentage points, mainly due to improvements on the single-principle problems. Although pre-to-post improvements were statistically significant ($F(1,39)=21.25$, $p<.0001$), no one group improved significantly more than any other group. This suggests that, at least for treatments lasting a short period of time, the improvement on problem solving was primarily due to practice in problem solving in general, not to any specific treatment.

Table 2: Percent Correct (S.D.) in Pre-, and Post-Problem-Test

Group	Pre-Test	Post-Test
HAT	29.4 (20.1)	41.3 (17.5)
EST	36.4 (25.8)	44.9 (25.9)
T	31.6 (24.6)	44.4 (24.4)

Given that the improvement of the HAT-group was not significantly better than that of the two control groups, we might ask whether the HAT-group was able to use the HAT appropriately? Our data indicate the answer is no: an analysis of the HAT-group's key-stroke data indicates that subjects were able to carry out appropriate analyses using the HAT on less than half of the treatment problems. Thus, the full potential of the HAT approach for improving problem solving skills cannot be fully evaluated until we ensure that subjects adopt the approach incorporated in the HAT.

GENERAL DISCUSSION

The results of Experiment 1 indicate that the hierarchical approach to problem solving, as exemplified by the HAT, helps students to shift their decision making criteria for problem categorization from one based on surface features toward one based on deep structure. We speculate that use of the HAT promotes this shift because it highlights the importance of applying principles to solve problems by asking subjects to select the applicable principle in the first menu they encounter. Even if the user is not able to answer all of the subsequent questions in the analysis correctly, the principle to be applied to obtain a solution may still be recognized as primary.

However, the current implementation of the hierarchical approach did not lead to significantly more improvement in problem solving than a "homework style" approach. Assuming the hierarchical approach is a potentially powerful tool for improving problem solving, there are two main reasons why the HAT treatment failed to yield more significant improvements in problem solving. First, subjects were not able to internalize the approach implicit in the HAT, since they had no way to gauge whether or not they were using the HAT appropriately. It appears that feedback and coaching are necessary ingredients to help novices assimilate the HAT's expert-like approach (Collins, Seely Brown & Newman, in press). Second, the treatment was relatively short, and therefore it is unrealistic to expect dramatic reorganization of declarative and procedural knowledge after using the HAT for only 5 hours.

Two implications appear to stand out. First, hierarchically structuring the problem analysis activities of novices holds promise for promoting expert-like behavior among novices. However, simply partaking in expert-like problem solving activities is not sufficient to promote dramatic improvements in a complex task such as problem solving--we need to guarantee that novices actually adopt the hierarchical approach before we can evaluate its full potential. Second, problem solving assessments may not be the most sensitive for measuring modest shifts toward expertise. Other measures, such as problem categorization, appear to be more sensitive assessments of subtle shifts toward expertise.

REFERENCES

Chase, W.G. & Simon, H.A. (1973). Perception in chess. Cognitive Psychology, 4, 55-81.

Chi, M.T.H., Feltovich, P.J. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Collins, A., Brown, J.S. & Newman, S.E. (in press). Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics. In L. Resnick (Ed.), Cognition and Instruction: Issues and Agendas. Hillsdale, NJ: Lawrence Erlbaum Assoc.

Dufresne, R., Gerace, W., Hardiman, P. & Mestre, J. (1987). Hierarchically structured problem solving in elementary mechanics: Guiding novices' problem analysis. Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics (Vol. III, pp. 116-130). Cornell University, Department of Education: Ithaca, NY.

Egan, D.E. & Schwartz, B.J. (1979). Chunking in recall of symbolic drawings. Memory & Cognition, 7, 149-158.

Ehrlich, K. & Soloway, E. (1982). An empirical investigation of the tacit plan knowledge in programming. Research Report #236, Department of Computer Science, Yale University.

Eylon, B.S. & Reif, F. (1984). Effect of knowledge organization on task performance. Cognition & Instruction, 1, 5-44.

Hardiman, P.T., Dufresne, R. & Mestre, J. (1987). Physics novices' judgments of solution similarity: When are they based on principles? Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics (Vol III, pp. 194-202). Cornell University, Department of Education: Ithaca, NY.

Heller, J.I. & Reif, F. (1984). Prescribing effective human problem solving processes: Problem description is physics. Cognition & Instruction, 1, 177-216.

Larkin, J. H. (1979). Information processing models in science instruction. In J. Lochhead & J. Clement (Eds.), Cognitive Process Instruction. Hillsdale, NJ: Lawrence Erlbaum Assoc.

Larkin, J.H., McDermott, J., Simon, D. & Simon, H. (1980). Models of competence in solving physics problems. Cognitive Science, 4, 317-345.

Mestre, J.P. & Gerace, W.J. (1986). Studying the problem solving behavior of experts and novices in physics via computer-based problem-analysis environments. Program of the Eighth Annual Conference of the Cognitive Science Society (pp. 741-746). Hillsdale, NJ: Lawrence Erlbaum Assoc.